SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1997	Technical - 97-25

4. TITLE AND SUBTITLE
Estimating the Number of Sinusoids and Its Performance Analysis

5. FUNDING NUMBERS
DAAH04-96-1-0082

6. AUTHOR(S)
Debasis Kundu

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Center for Multivariate Analysis
Department of Statistics
417 Thomas Bldg.
Penn State University
University Park, PA 16802

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

19971215 080

13. ABSTRACT (Maximum 200 words)
Detecting the number of signals and estimating the parameters of the signals is an important problem in signal processing. Quite a number of papers appeared in the last twenty years regarding the estimation of the parameters of the sinusoidal components but not that much of attention has been given in estimating the number of terms presents in a sinusoidal signal. Fuchs developed a criterion based on the perturbation analysis of the data auto correlation matrix to estimate the number of sinusoids, which is in some sense a subjective-based method. Recently Reddy and Biradar proposed two criteria based on AIC and MDL and developed an analytical framework for analyzing the performance of these criteria. In this paper we develop a method using the extended order modelling and singular value decomposition technique similar to that of Reddy and Biradar. We use penalty function technique but instead of using any fixed penalty function like AIC or MDL, a class of penalty functions satisfying some special properties has been used.

14. SUBJECT TERMS
Sinusoidal Signals, Law of Iterated Logarithm, Strong consistency, Monte Carlo Simulation

15. NUMBER IF PAGES
14

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

ESTIMATING THE NUMBER OF SINUSOIDS AND ITS PERFORMANCE ANALYSIS

Debasis Kundu

Technical Report 97-25

December 1997

Center for Multivariate Analysis
417 Thomas Building
Penn State University
University Park, PA 16802

Research work of author was supported by the Army Research Office under Grant DAAHO4-96-1-0082.

Estimating the Number of Sinusoids and Its Performance Analysis

Debasis Kundu
Department of Mathematics
Indian Institute of Technology Kanpur
Pin 208016, India

Abstract: Detecting the number of signals and estimating the parameters of the signals is an important problem in signal processing. Quite a number of papers appeared in the last twenty years regarding the estimation of the parameters of the sinusoidal components but not that much of attention has been given in estimating the number of terms presents in a sinusoidal signal. Fuchs developed a criterion based on the perturbation analysis of the data auto correlation matrix to estimate the number of sinusoids, which is in some sense a subjective-based method. Recently Reddy and Biradar proposed two criteria based on AIC and MDL and developed an analytical framework for analyzing the performance of these criteria. In this paper we develop a method using the extended order modelling and singular value decomposition technique similar to that of Reddy and Biradar. We use penalty function technique but instead of using any fixed penalty function like AIC or MDL, a class of penalty functions satisfying some special properties has been used. We prove that any penalty function from that special class will give consistent estimate under the assumptions that the error random variables are independent and identically distributed with mean zero and finite varinace. We also obtain the probabilities of wrong detection for any particular penalty function under somewhat weaker assumptions than that of Reddy and Biradar or Kaveh et al.. It gives some idea to choose the proper penalty function for any particular model. Simulations are performed to verify the usefulness of the analysis and to compare our methods with the existing ones.

Key Words and Phrases: Sinusoidal Signals, Law of Iterated Logarithm, Storng consistency, Monte Carlo Simulation

AMS Subject Classifications(1985): 62J02, 65D10

Short Running Title: Estimating Number of Sinusoids

e-mail: kundu@iitk.ernet.in

1 Introduction

The problem of detecting the number of terms and estimating the parameters of sinuoids in presence of noise is an important problem in signal processing. In the last twenty years several methods have been proposed to estimate the frquencies of the sinusoids very efficiently, see [1], but not much attention has been paid in estimating the number of sinusoids presents in any particular signal. Tufts and Kumaresan [7] proposed the modified forward backward linear prediction (MFBLP) technique, which is capable of estimating the poorly separated frquencies, with short data lengts and moderate SNR, but unfortunately it is known to give inconsistent estimates (see [10] and [13]). They have also given a graphical technique to estimate the number of sinusoids, which is very subjective. Some of the other techniques which are available in the literature for example [8], [9], [3] can be used, but they also depend on the subjective choice of the individual and therefore makes it difficult to implement in practice.

Rao [10] proposed an information theoretic criterion to estimate the number of signals for undamped exponential model but it is observed [5] that Rao's suggestion is very dufficult to implement in practice. A practical implementation procedure has been suggested in [5] but it is observed that the proposed method depends very much on the penalty function used. Some suggestions about the penalty function have been given mainly based on the computer simulation but not on any proper analysis. Under the assumptions of normality of the error components, recently Reddy and Biradar [2] proposed two criteria one is AIC type and the other one is MDL type based on the singular value decomposition technique and using both the forward and backward linear predictor equation. They also obtained the probability of over as well as under estimation following the assumption made by [4] and [11]. They use the assumptions of normality of the error random variables in developing the criteria as well as in the performance analysis. The performance looks quite satisfactory.

In this paper, we develop a method mainly based on the extended order modelling and singular value decomposition technique. We use penalty functions similarly as AIC and MDL but instead of any particular penalty function a class of penalty functions satisfying some special properties has been used. We prove that any penalty function from that particular class will give consistent estimate. We only assume that the errors are independent and identically distributed (i.i.d.) with mean zero and finite variance to prove the consistency result. In fact we do not need any distributioinal assumptions on the error random variables. Next we obtain an estimate on the probability of wrong detection under the assumptions that the errors are i.i.d. Gaussian random variables with mean zero and finite variance. We obtain the probabilty of wrong detection using matrix perturbation technique and large sample approximation. Once we obtain the probability of wrong detection for any particular penalty function, we choose that penalty function for which the probability of wrong detection is minimum. Some simulations are performed to see the usefulness of our method and to compare it with the other existing ones.

The rest of the paper is organised as follows. In Section 2, we give the estimation procedure and the consistency result is proved in Section 3. The performance analysis is carried out in Section 4 and some numerical experiments are performed in Section 5. The choice of the penalty function is suggested in Section 6 and finally we draw the conclusions of our work in Section 7.

2 Estimation Procedure

Consider N uniformly spaced data samples $y(n)$ of M real sinusoids corrupted by additive white noise

$$y(n) = \sum_{i=1}^{M} a_i sin(2\pi\omega_i n + \phi_i) + \epsilon(n); n = 1, ..., N \tag{1}$$

where a_i's are the amplitudes, can be any arbitrary real numbers, $\omega_i's$ are the angular frequencies lying between 0 and .5, $\phi_i's$ are the initial phases of the i th frequency lying between 0 to 2π. ϵ(n)'s are i.i.d. random variables with mean zero and finite variance σ^2. M, the number of signals, is assumed to be unknown. Given a sample of size N, the problem is to estimate M. Let us assume that the number of sinusoids can be at most K, which is known, i.e. $M \leq K$. Consider the following data matrices;

$$\mathbf{A_F} = \begin{pmatrix} y(1) & \dots & y(L) \\ \vdots & \vdots & \vdots \\ y(N-L+1) & \dots & y(N) \end{pmatrix}, \tag{2}$$

and

$$\mathbf{R_N} = \frac{1}{(N-L+1)}[\mathbf{A_F^T A_F}] \tag{3}$$

here L is any integer such that $2K < L < N - 2K + 1$. Let the eigenvalues of $\mathbf{R_N}$ be $\hat{\lambda}_1 >, \dots, > \hat{\lambda}_L$. Compute

$$IC(k) = \log(\hat{\lambda}_{2k+1} + 1) + kC_{N-L+1} \tag{4}$$

for k = 1, ..., $2K$, here C_N (penalty functions) satisfies the following conditions

$$(a) C_N > 0 \qquad (b) C_N \to 0 \qquad (c) \frac{C_N \sqrt{N}}{\sqrt{loglogN}} \to \infty \tag{5}$$

Then our criterion is the following; Choose $\hat{M}$ such that

$$IC(\hat{M}) = min[IC(1), \ldots, IC(K)] \tag{6}$$

So $\hat{M}$ is an estimate of M. In the next section we prove that $\hat{M}$ converges to M almost surely for any C_N, satisfying (5), under the assumptions that the errors are i.i.d. with mean 0 and finite varinace $\sigma^2 > 0$.

3 Consistency Results

Let us denote the (p, q)th element of the matrix $\mathbf{A_F^T A_F}$ by a_{pq}. Therefore

$$a_{pq} = \sum_{s=0}^{N-L} y(p+s)y(q+s) \tag{7}$$

Now write;

$$\begin{aligned} y(n) &= \sum_{i=1}^{M} a_i sin(2\pi\omega_i n + \phi_i) + \epsilon(n) \\ &= \sum_{i=1}^{M} \frac{a_i}{2j}[exp(j2\pi\omega_i n)exp(j\phi_i) - exp(-j2\pi\omega_i n)exp(-j\phi_i)] + \epsilon(n) \\ &= \sum_{i=1}^{2M} A_i exp(j2\pi\eta_i n) + \epsilon(n) \end{aligned} \tag{8}$$

where $j = \sqrt{-1}$, $A_1 = \frac{a_1}{2j}\, e^{j\phi_1}$, $A_2 = -\frac{a_1}{2j}\, e^{-j\phi_1}, \eta_1 = \omega_1, \eta_2 = -\omega_1$, and the others are defined similarly. Since $y(q+s) = \bar{y}(q+s)$ ('-' denotes the complex conjugate),

$$\begin{aligned} a_{pq} &= \sum_{s=0}^{N-L} y(p+s)\bar{y}(q+s) \\ &= \sum_{s=0}^{N-L} (\sum_{i=1}^{2M} A_i exp(j2\pi\eta_i(p+s)) + \epsilon(p+s)) \times \\ &\quad (\sum_{i=1}^{2M} \bar{A}_i exp(-j2\pi\eta_i(q+s)) + \bar{\epsilon}(q+s)) \\ &= \sum_{s=0}^{N-L} (\sum_{i=1}^{2M} A_i exp(j2\pi\eta_i(p+s)))(\sum_{i=1}^{2M} \bar{A}_i exp(-j2\pi\eta_i(q+s))) \end{aligned}$$

$$+\sum_{s=0}^{N-L}\epsilon(p+s)(\sum_{i=1}^{2M}\bar{A}_i exp(-j2\pi\eta_i(q+s)))$$
$$+\sum_{s=0}^{N-L}\bar{\epsilon}(q+s)(\sum_{i=1}^{2M}A_i exp(j2\pi\eta_i(p+s)))$$
$$+\sum_{s=0}^{N-L}\epsilon(p+s)\bar{\epsilon}(q+s) \quad (9)$$

Therefore by the law of the iterated logarithm of the $M-dependent$ sequence, we can say that for fixed L as N tends to infinity, we have

$$\mathbf{R_N} = \frac{1}{N-L+1}\mathbf{A_F^T A_F} = \Omega\mathbf{D}\Omega^H + \sigma^2 I + O(\frac{loglog(N-L+1)}{N-L+1})^{1/2} \; (a.s.) \quad (10)$$

here 'a.s.' stands for almost surely and the matrix Ω and $\mathbf{D}$ are as follows:

$$\Omega = \begin{pmatrix} exp(2\pi\eta_1) & \dots & exp(2\pi\eta_{2M}) \\ \vdots & \vdots & \vdots \\ exp(2\pi\eta_1 L) & \dots & exp(2\pi\eta_{2M} L) \end{pmatrix}, \mathbf{D} = \mathrm{diag}(\| A_1\|^2, \dots, \| A_{2M}\|^2) \quad (11)$$

Let us denote the matrix $\Omega\mathbf{D}\Omega^H$ by Σ and let the non zero eigenvalues of Σ be $\mu_1 > \dots > \mu_{2M}$. Let us denote the ordered eigenvalues of $\Sigma + \sigma^2\mathbf{I}$ by $\lambda_1 > \dots \lambda_{2M} > \lambda_{2M+1} = \dots = \lambda_L = \sigma^2$, $\lambda_i = \mu_i + \sigma^2$ for $i = 1, \dots, 2M$ and $\lambda_i = \sigma^2$ for $i = 2M+1, \dots, L$. We need the following result for further development:

Lemma 1 : Let $\mathbf{P}$ and $\mathbf{Q}$ be two Hermitian matrices of order $m \times m$ and let the spectral decomposition of $\mathbf{P}$ and $\mathbf{Q}$ be as follows

$$\mathbf{P} = \sum_{\mathbf{i=1}}^{\mathbf{m}} \delta_\mathbf{i}\mathbf{p_i p_i^H} \qquad \mathbf{and} \qquad \mathbf{Q} = \sum_{\mathbf{i=1}}^{\mathbf{m}} \gamma_\mathbf{i}\mathbf{q_i q_i^H} \quad (12)$$

where $\delta_1 \leq \dots \leq \delta_m$ and $\gamma_1 \leq \dots \leq \gamma_m$ are the ordered eigenvalues and $\mathbf{p_i}$'s, $\mathbf{q_i}$'s are the corresponding orthonormal eigenvectors of $\mathbf{P}$ and $\mathbf{Q}$ respectively, then if $|p_{ik} - q_{ik}| \leq \alpha$ for all $i, k = 1, \dots, m$ and for some α, then there is a constant C such that $|\delta_i - \gamma_i| \leq C\alpha$.

Proof : It mainly follows from von-Neumann's inequality but see [12] for details.

Using Lemma 1 and (10) we obtain

$$\hat{\lambda}_i = \lambda_i + O(\frac{loglog(N-L+1)}{N-L+1})^{1/2} \; (a.s.) \quad (13)$$

Now to prove $\hat{M}$ is a consistent estimate of M observe that it is enough to prove

$$IC(q,\, C_{N-L+1}) - IC(M,\, C_{N-L+1}) > 0; \qquad for \qquad q = 1, \dots M-1, M+1, \dots K \quad (14)$$

as $N \to \infty$.

Case I: $q < M$

$$\begin{aligned} IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) &= \log(\hat{\lambda}_{2q+1}) - \log(\hat{\lambda}_{2M+1}) + (q-M)\ C_{N-L+1} \\ &\to \log(\lambda_{2q+1}+1) - \log(\lambda_{2M+1}+1) \\ &= \log(\mu_{2M+1}+\sigma^2+1) - \log(\sigma^2+1) > 0 \quad (15) \end{aligned}$$

Case II: $q > M$

$$\begin{aligned} IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) &= \log(\hat{\lambda}_{2q+1}) - \log(\hat{\lambda}_{2M+1}) + (q-M)\ C_{N-L+1} \\ &\to \log(1+\sigma^2+\ h_1) - \log(1+\sigma^2+\ h_2) \\ &\quad +(q-M)\ C_{N-L+1} \\ &= \log(1+\sigma^2) +\ h_1\frac{1}{1+\sigma^2} -\ h_2\frac{1}{1+\sigma^2} \\ &\quad -\log(1+\sigma^2) + (q-M)\ C_{N-L+1} \\ &\quad +O(\frac{loglog(N-L+1)}{N-L+1}) \qquad (16) \end{aligned}$$

here $h_1 = O(\frac{loglog(N-L+1)}{N-L+1})^{1/2} =\ h_2$.
Now observe that;

$$\begin{aligned} (IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1})) &= (q-M) + \\ &\quad \frac{1}{C_{N-L+1}}O(\frac{\log\log(N-L+1)}{N-L+1})^{1/2} \\ &\quad +\frac{1}{C_{N-L+1}}O(\frac{loglog(N-L+1)}{N-L+1}) \quad (17) \end{aligned}$$

From the properties of C_N (see (5)), it follows that the second and the third term of the right hand side of (17) goes to zero as N tends to infinity, therefore (17) implies

$$\frac{1}{C_{N-L+1}}(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1})) \to (q-M) > 0 \qquad (18)$$

Therefore we can conclude that for large $(N-L+1)$

$$(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1})) > 0 \qquad (19)$$

Combining (15) and (19) we obtain (14).

4 Performance Analysis

In this section we obtain an upper bound for $P(\hat{M} \neq M)$. Now

$$P(\hat{M} \neq M) \ = \ P(\hat{M} < M) + P(\hat{M} > M)$$

$$
\begin{aligned}
&= \sum_{q=0}^{M-1} P(\hat{M}=q) + \sum_{q=M+1}^{K} P(\hat{M}=q) \\
&= \sum_{q=0}^{M-1} P(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) < 0) \\
&\quad + \sum_{q=M+1}^{K} P(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) < 0) \qquad (20)
\end{aligned}
$$

Let's consider two different cases:
Case I, $q < M$

$$
\begin{aligned}
& P(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) < 0) \\
= \ & P(log(\hat{\lambda}_{2q+1}+1) - log(\hat{\lambda}_{2M+1}+1) + (q-M)\ C_{N-L+1} < 0) \\
= \ & P(log(\lambda_{2q+1}+1) - log(\lambda_{2M+1}+1) + (q-M)\ C_{N-L+1} < \\
& log(\hat{\lambda}_{2M+1}+1) - log(\lambda_{2M+1}+1) + log(\lambda_{2q+1}+1) - log(\hat{\lambda}_{2q+1}+1)) \\
< \ & P(log(\lambda_{2q+1}+1) - log(\lambda_{2M+1}+1) < (M-q)\ C_{N-L+1} + \\
& |log(\hat{\lambda}_{2M+1}+1) - log(\lambda_{2M+1}+1)| + |log(\hat{\lambda}_{2q+1}+1) - log(\lambda_{2q+1}+1)|) \quad (21)
\end{aligned}
$$

Let δ be such that for large N

$$
log(\lambda_{2q+1}+1) - log(\lambda_{2M+1}+1) > (M-q)\ C_{N-L+1} + \delta \qquad (22)
$$

Therefore for large N

$$
P(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) < 0) = 0 \qquad (23)
$$

Case II, $q > M$

$$
\begin{aligned}
& P(IC(q,\ C_{N-L+1}) - IC(M,\ C_{N-L+1}) < 0) \\
= \ & P(\log(\hat{\lambda}_{2q+1}+1) - \log(\hat{\lambda}_{2M+1}) + (q-M)\ C_{N-L+1} < 0) \\
= \ & P(\log(\hat{\lambda}_{2M+1}+1) - \log(\hat{\lambda}_{2q+1}+1) > (q-M)\ C_{N-L+1}) \qquad (24)
\end{aligned}
$$

Therefore observe that from (23) it is clear that for large N the probability of underestimation is zero and to obtain (24) we need to know the joint distributions $\hat{\lambda}_k$ for $k = 2M+1, \ldots, L$. Let's assume at this point that ϵ(n)'s are i.i.d. Gaussian random variables with mean zero and variance σ^2. Later on we see that it is possible to relax this assumption. Let's denote the matrix $(\Sigma + \sigma^2 \mathbf{I})$ by $\mathbf{R}$. Therefore from the central limit theorem, we can say that asymptotically

$$
\sqrt{N-L+1}(Vec(\mathbf{R_{N-L+1}}) - Vec(\mathbf{R})) \qquad (25)
$$

will be normally distributed with mean vector zero and certain $L^2 \times L^2$ variance covariance matrix Γ (say). Here Vec(.) of an $L \times L$ matrix denotes the $L^2 \times 1$ vector stacking the columns one below the other. Using the perturbation theory ([15]; page 66) let's write

$$\mathbf{R_{N-L+1}} = \mathbf{R} + (\mathbf{R_{N-L+1}} - \mathbf{R}) = \mathbf{R} + \epsilon\mathbf{B} \tag{26}$$

Here $0 < \epsilon \ll 1$, and $\mathbf{B}$ is a Hermitian, zero mean matrix with elements that are asymptotically jointly Gaussian, which follows from (25). Let λ_i be any particular eigenvalue of $\mathbf{R}$ and let the corresponding normalized eigenvector be $\mathbf{z_i}$. Observe that if λ_i is a multiple eigenvalue of $\mathbf{R}$, then $\mathbf{z_i}$ is not unique, but then we take one particular $\mathbf{z_i}$ (see [15] page 69). Let $\hat{\lambda}_i$ be the corresponding perturbed eigenvalue of $\mathbf{R_{N-L+1}}$, then from ([15] page 69) it follows that

$$\hat{\lambda}_i = \lambda_i + \epsilon(\mathbf{z_i^T B z_i}) \tag{27}$$

Since the elements of $\mathbf{B}$ are normally distributed, therefore $\mathbf{z_i^T B z_i}$, which is a linear combination of the elements of $\mathbf{B}$, will also be normally distributed with mean zero and finite varinace. Now from (27) it follows that

$$E(\hat{\lambda}_i) \approx \lambda_i \tag{28}$$

Now to compute the covariance between $\hat{\lambda}_i$ and $\hat{\lambda}_j$, let $\mathbf{z_i}$ and $\mathbf{z_j}$ be any two orthonormal eigenvectors corresponding to λ_i and λ_j respectively, therefore we have

$$\hat{\lambda}_i = \lambda_i + \epsilon(\mathbf{z_i^T B z_i}) \tag{29}$$

$$\hat{\lambda}_j = \lambda_j + \epsilon(\mathbf{z_j^T B z_j}) \tag{30}$$

Therefore

$$E(\hat{\lambda}_i - \lambda_i)(\hat{\lambda}_j - \lambda_j) = \epsilon^2 E(\mathbf{z_i^T B z_i})(\mathbf{z_j^T B z_j}) \tag{31}$$

Since we are mainly interested about the distribution of the repeated eigenvalues, so let's take $\mathbf{z_i}$ and $\mathbf{z_j}$ be any two orthonormal eigenvectors corresponding to σ^2, therefore we can obtain after some simplifications;

$$E(\mathbf{z}_i^T \mathbf{B} \mathbf{z}_i)(\mathbf{z}_j^T \mathbf{B} \mathbf{z}_j) = \frac{\sigma^4}{(N-L+1)^2\epsilon^2} \sum_{p=1}^{N-L+1} \sum_{q=1}^{N-L+1} (\mathbf{z}_i^T \Gamma_{pq} \mathbf{z}_j)(\mathbf{z}_j^T \Gamma_{qp} \mathbf{z}_i) \tag{32}$$

Here Γ_{pq}'s are $L \times L$ matrices and they are defiend as follows $\Gamma_{pq} = \Gamma_{qp}^T$ and $\Gamma_{pq} = 0$ if $(q-p) > L$. For $q > p, \Gamma_{pq}$ has all the elements to be zero except at the positions $(q-p+1, 1), \ldots, (L, p-q+L)$ which are ones, as follows;

$$\Gamma_{pq} = \begin{pmatrix} 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 1 & 0 \end{pmatrix}, \tag{33}$$

Therefore we obtain that $\hat{\lambda}_{2M+1}$ and $\hat{\lambda}_{2q+1}$ for $q > M$ are jointly normal each of them with mean σ^2 and it has the variance covarinace matrix which can be obtained from (32).

Observe that even if we know the joint distribution of $\hat{\lambda}_i$ for $i = 2M+1, \ldots, L$, theoretically it is very difficult to compute (24). We use simulation technique similarly as [11] to compute (24) by using the joint distribution $\hat{\lambda}_{2M+1}$ and $\hat{\lambda}_{2q+1}$. The details will be explained in the next section.

Observe that since we assume that the eigenvalues of Σ are distinct therefore the distribution of $\hat{\lambda}_1, \ldots, \hat{\lambda}_{2M}$ will be independent of $\hat{\lambda}_{2M+1}, \ldots, \hat{\lambda}_L$ asymptotically. Moreover they will be jointly normal with mean λ_i and variance $\frac{\lambda_i^2}{N-L+1}$ for $i = 1, \ldots, 2M$ see [14]. and they will be independent of each other.

5 Numerical Experiments

In this section we perform some numerical experiments to present both the effectiveness of our method and the usefulness of the analysis. We consider the same model as that of [2].

The data sample are generated from the following model:

$$y(n) = \sqrt{20} sin(2\pi\omega_1 n) + \sqrt{20} sin(2\pi\omega_2 n + \phi) + \epsilon(n); \qquad n = 1, \ldots, N. \tag{34}$$

here $\omega_1 = .2$ and $\omega_2 = .2 + \delta$ with $\delta = \frac{1}{N}$ and $N = 64$. Here ϵ(n)'s are i.i.d. Gaussian random variables with mean zero and variance σ^2, which is chosen appropriately to give the desired signal to noise ratio (SNR) defined as SNR $= 10 log_{10} 10/\sigma^2$. We use twelve different C_N, all of them satisfying (5) but converging to zero at different rates. We define them as $C_N^{(1)}, \ldots, C_N^{(12)}$. They are as follows; $C_N^{(1)} = (\frac{1}{N})^{.1}, C_N^{(2)} = (\frac{1}{N})^{.2}, C_N^{(3)} = (\frac{1}{N})^{.3}, C_N^{(4)} = (\frac{1}{N})^{.4}, C_N^{(5)} = \frac{1}{\log N}, C_N^{(6)} = (\frac{1}{\log N})^{\cdot}2, C_N^{(7)} = (\frac{1}{\log N})^{\cdot}4$, $C_N^{(8)} = (\frac{1}{\log N})^{\cdot}6, C_N^{(9)} = (\frac{1}{\log N})^{\cdot}8, C_N^{(10)} = (\frac{1}{N \log N})^{\cdot}1, C_N^{(11)} = (\frac{1}{N \log N})^{\cdot}3, C_N^{(12)} = (\frac{1}{\log \log N})$. We take $K = 8$, i.e. the maximum number of sinusoids, same as [2] and $L = 32$. Out of 500 simulations, the percentage of correct estimate (PCE), the percentage of under estimate(PUE) and the percentage of over estimate (POE) are reported for SNR = 5dB and SNR = 10dB. We also obtain the theoretical value for the upper bound of the probability of over estimate as follows. We draw a sample of size (L-2M) from Gaussian random variable with mean σ^2 and variance covariance matrix given by (32), we order them as $\hat{\lambda}_i$ for $i = 2M, \ldots, L$ and check whether $(log\hat{\lambda}_{2M+1} - log\hat{\lambda}_{2q+1}) > (q - M)\ C_{N-L+1}$ is true or not for $q = M+1, \ldots, K$. We repeat this process 5000 times and compute the percentage of times it is true and that gives an estimate of (24). Observe that although we have seen that the probability of underestimate will be zero for large sample size, we have calculated the probability of underestimate using the distribution properties of the $\hat{\lambda}_1, \ldots, \hat{\lambda}_{2M+1}$ and using the same procedure as above by repeating over 5000 times. The results are reported in Table 1 for SNR = 5dB and in Table 2 for SNR = 10dB. The quantity within the bracket indicate the theoretical bounds of POE's and PUE's.

Table 1

$C_N^{(i)}$	PUE	PCE	POE
$C_N^{(1)}$	0(0)	97	3(5)
$C_N^{(2)}$	0(0)	84	16(22)
$C_N^{(3)}$	0(0)	59	41(50)
$C_N^{(4)}$	0(0)	31	69(82)
$C_N^{(5)}$	0(0)	49	51(56)
$C_N^{(6)}$	0(0)	100	0(3)
$C_N^{(7)}$	0(0)	94	6(14)
$C_N^{(8)}$	0(0)	82	18(23)
$C_N^{(9)}$	0(0)	64	36(39)
$C_N^{(10)}$	0(0)	94	6(12)
$C_N^{(11)}$	0(0)	29	71(82)
$C_N^{(12)}$	0(0)	98	2(5)

Table 2

$C_N^{(i)}$	PUE	PCE	POE
$C_N^{(1)}$	0(0)	100	0(0)
$C_N^{(2)}$	0(0)	92	8(8)
$C_N^{(3)}$	0(0)	66	34(28)
$C_N^{(4)}$	0(0)	46	54(60)
$C_N^{(5)}$	0(0)	55	45(41)
$C_N^{(6)}$	0(0)	100	0(0)
$C_N^{(7)}$	0(0)	96	4(2)
$C_N^{(8)}$	0(0)	90	10(9)
$C_N^{(9)}$	0(0)	75	25(23)
$C_N^{(10)}$	0(0)	96	4(2)
$C_N^{(11)}$	0(0)	45	55(60)
$C_N^{(12)}$	0(0)	100	0(0)

At finite sample size the performance of the proposed method very much depends on the penalty function used, although all of them give consistent estimates as the sample size tends to infinity. From Table 1 and Table 2 it is clear that the performance of all the methods becomes worse at low SNR, which is not very surprising. It is important to observe that the theoretical probability matches quite well in almost all cases considered and the estimates are better in most of the cases at high SNR. It seems that the approximations made for the large sample work reasonably well at moderate sample sizes and for mederate SNR.

6 How to Choose the Penalty Function?

Looking at the tables it is clear that the theoretical bounds are quite close to the actual one. But unfortunately without knowing the original parameters we can not calculate the theoretical probabilites. Reddy and Biradar [2] also did not raise this question that how to estimate these bounds. One way of course by replacing these parameters values by their estimates. We would like to estimate these probabilities with the help of the given sample and use it to choose the proper penalty function. The idea is as follows: From any particular realization of the model, we compute the matrix $\mathbf{R_{N-L+1}}$ and obtain the eigenvalues and the corresponding eigenvectors. Now suppose using the penalty function $C_N^{(k)}$, we estimate the order of the model as M_k. Assuming M_k is the correct order model we compute the estimate of σ^2 by averaging the last $L - 2\ M_k$ eigenvalues. Now using the joint distribution $\hat{\lambda}_i$ for $i = 2\ M_k + 1, \ldots, L$, (24) can be calculated by using the simulation technique as discussed in the previous section. Similarly from $\hat{\lambda}_1, \ldots, \hat{\lambda}_{2M_k+1}$, we can compute an estimate of the probability of under estimate. Therefore adding the two, we can obtain an estimate of the upper bound of the probability of wrong detection. It can be shown easily that for large N, the estimate of the probabiltiy of wrong detection under the asuumption of correct order model will be less than the estimate of wrong detection under the assumption of lower/higher order model, because the former one goes to zero as N tends to infinity, where as the later one goes to a positive quantity.

We use this idea and compute the estimate of the probabiltiy of wrong detection for all the criteria and choose that one which gives the lowest estimate of the probability of wrong detection. We have used the same model and the same set of penalty functions and in each trial we choose that penalty function which gives the lowest estimate of the probabilty of wrong detection. In each trial we run 100 simulation to compute the estimate of the probability of error and it is repeated over 500 trials. The result is reported in Table 3, which indicates the percentage of underestimate, correct estimate and over estimate out of these 500 replications.

Table 3

SNR	PUE	PCE	POE
5dB	0	94	6
10dB	0	95	5

From Table 3, it is observed that the proposed method works quite well. For the same model it is observed [2] that when SNR = 5dB, the simulation shows that MDL criterion can detect at most 88 percent and at SNR = 10dB it can detect at most 91 percent correctly but in our case it is observed that it can detect 94 percent and 95 percent respectively. It is higher than what they have obtained, which is not very surprising because we have a class of penalty functions and from there we are trying to choose the best.

Another question naturally comes that how to choose the class of penalty function. We suggest the following way; take any particular class of resonable size may be around ten or twelve then obtain the estimate of the probability of wrong detection for all the methods and compute the minimum value, if the minimum itself is high that means the error has been calculated using wrong order model and the class of penalty function is not good and on the other hand if it is low, that means this class is fine. We have computed the mean and standard deviation of the minimum probabiltiy of error when the model has been chosen correctly and when the model has not been chosen correctly for the same experiments. The results are reported below in Table 4a and Table 4b, where Table 4a gives the mean and standard deviation (s.d) of the minimum probability of error when the model has been chosen correctly and Table 4b gives the result when the model has not been chosen correctly.

Table4a

SNR	mean	s.d
5dB	.0489	.0845
10dB	.0255	.0527

Table4b

SNR	mean	s.d
5dB	.2850	.1360
10dB	.2540	.1288

From the Tables 4a and 4b it is clear that the mean of the minimum probability of error under the assumptions of the correct order model is much smaller than under the assumption of wrong order model. So it is expected that the minimum probabiltiy of error can give us some idea whether the class is sufficient or not.

7 Conclusions

In this paper, we propose a method to detect the number of sinusoids present in a sinusoidal signal corrupted by additive noise. We use extended order modelling and singular value decomposition technique. We prove the strong consistency of the proposed method under the assumptions that the errors are i.i.d. with mean zero and finite variance. We have done some performance analysis of the proposed method under the assumption that the errors are Gaussian. But theoretically speaking it is possible to relax this condition. Observe that (25) is true asymptotically even if the errors are not Gaussian. But in that case more sample size is required. Another important point is to observe that we can use both the backward and forward data to compute $\mathbf{R_N}$ of Section 2 and in that case it is not very difficult to prove that the proposed criteria will be consistent but unfortunately in that case the performance analysis becomes difficult, particularly the expression (32) will have more terms so it is not pursued here. It is observed that our method behaves quite satisfactory and works better than the existing methods.

Acknowledgements: Most of the work was done when the author was visiting the ECE Department of the Indian Institute of Science, Bangalore. The author would like to thank Professor V.U. Reddy for his encouragement and help to prepare this manuscript.

References

[1] P. Stoica, "List of References on Spectral Line Analysis", *Signal Processing*, 31, pp 329 - 340, 1993.

[2] V.U.Reddy and L.S.Biradar, "SVD-Based Information Theoretic Criteria for Detection of the Number of Damped/Undamped Sinusoids and Their Performance Anlysis", *IEEE Trans. on Signal Processing*, 41, 9, pp 2872-2881, 1993.

[3] J.J.Fuchs, "Estimating the Number of Sinusoids in Additive White Noise", *IEEE Trans. Acoust. Speech, Signal Processing*, 36, 12, pp 1846-1853, 1988.

[4] H.Wang and M. Kaveh, "On the Performance of Signal-Subspace Processing Part I: Narrow -Band Systems", *IEEE Trans. Acoust., Speech, Signal Processing*, 34,5, pp 1201-1209, 1986.

[5] D. Kundu, "Detecting the Number of Signals for Undamped Exponential Models Using Information Theoretic Criteria", *Journal of Statistical Computation and Simulation*, 44, pp 117-131, 1992.

[6] D.Kundu and R. Kundu, "Consistent Estimates of Super Imposed Signals when Some Observations are Missing", *Journal of Statistical Planning and Inference*, 44, pp 205-218, 1995.

[7] D.W.Tufts and R. Kumaresan, "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood", *Proc. IEEE*, 70,9, pp 975-989, 1982.

[8] M.S.Bartlett, "A Note on the Multiplying Factors for Various χ^2 Approximations", *Jour. Royal Stat. Soc. Ser. B*, 16, pp 296-298, 1954.

[9] D.N. Lawley, "Tests of Significance of the Latent Roots of the Covariance and Correlation Matrices", *Biometrica*, 43, pp 128-136, 1956.

[10] C.R.Rao, "Some Results in Signal Detection", *Stat. Dec. Theo. and Rel. Top., IV* Eds Gupta, S.S. and Berger, J., pp 319-332, Springer Verlag, New York, 1988.

[11] M.Kaveh, H. Wang and H. Hung, "On the Theoretical Performance of a Class of Estimators of the Number of Narrow-Band Sources", *IEEE Trans. Acoust. Speech, Signal Processing*, 35, 9, pp 1350-1352, 1987.

[12] Z.D.Bai and C.R.Rao, "Spectral Analytic Methods for the Estimation of Number of Signals and Directions of Arrivals", *Spectral Analysis in One or Two Dimensions*, eds. S. Prasad and L. Kashyap, Oxford and IBH Publications, New Delhi, pp 493-507, 1989.

[13] D.Kundu and A. Mitra, "Consistent Method of Estimating Superimposed Exponential Signals", *Scandinavian Journal of Statistics*, 22,1, pp 73-82, 1995.

[14] D.R.Brillinger, *Time Series: Data Analysis and Theory*, Expanded Edition, San Francisco, CA: Holden-Day, 1981

[15] W.H.Wilkinson, *The Algebraic Eigenvalue Problems*, Oxford University Press, 1965